Sports Info Solutions Football Analytics Challenge

Ryan Keeney
keeney.ryan@gmail.com

Hello

I’m Ryan, enjoy!

Table of Contents

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

Talent Distribution

Question 2 & 3

Limitations and Future Analysis

Key findings:

1. The further outside a player was located based on the Technique, the higher their true pressure rate.
2. Individual player’s true pressure rate correlates to a team's success rate[1] (adj. R-squared ~ 0.1-0.3) against the pass as well as a team’s actual pressure rate (adj. R-squared ~ 0.5-0.6) if we include additional factors[2].
3. This analysis focuses on pass plays, because I didn’t find any correlation to rush plays results with the data provided. It may exist, but I didn’t find anything meaningful.

---

(1) Success Rate = 1 if the play resulted in 0 > EPA, or 0 if the play resulted in < 0 EPA.

(2) Such as a teams blitz rate and the average defenders on the line per play

- Key Findings
- Process & Assumptions
- General Findings
- True Pressure Rate
- Positional Value
    - Question 1
- Player Talent
- Talent Distribution
    - Question 2 & 3
- Limitations and Future Analysis

Process and Assumptions

The data only required a few updates for my analysis.

a. Removed 1x play with NA EPA. GameID: 2757, EventID: 915
b. Added play “success” category (EPA > 0 = 1, else = 0)
c. Added of defenders recorded, number of DL and LB players listed by “OnFieldPosition”, and the number of rushers
d. Added “Blitz” category, for when the number of rushers recorded is > 4 players.
e. Added “shotgun” play description attribute.
f. Updated EventID based on the play description provide to remove broken plays, kneels, spikes and sort QB scrambles (pooled with pass plays with a tag) vs. QB designed runs (pooled with run plays).

Process and Assumptions

Initial play-by-play data assumptions

a. I did not remove “scramble” plays from the passing plays since they include if a defender was rushing, even though none of these plays recorded a “PressureOnPlay”.

Process and Assumptions

Summary

a. Data analyzed in R studio, using logistic regressions and ANOVA+Tukey's range test depending on the data types.
b. Analyzed what correlated to pass and rush success on a team level, as well as secondary factors such as pressure, tackle rates, intended
c. Selected attributes to focus on (pressure, pass success rate)
d. Evaluated play-by-play positional data to see what correlated to these metrics (Question 1)

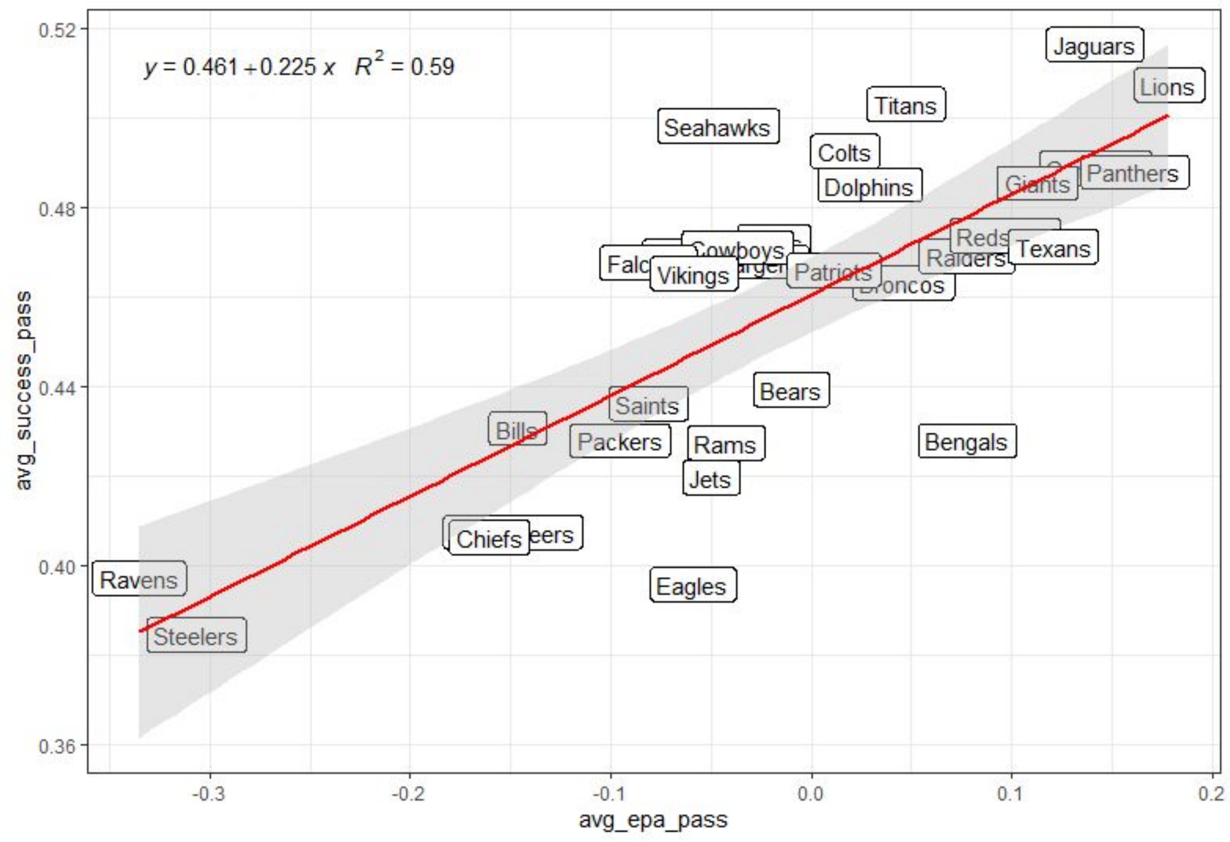
Findings

First, I looked at the data by team.

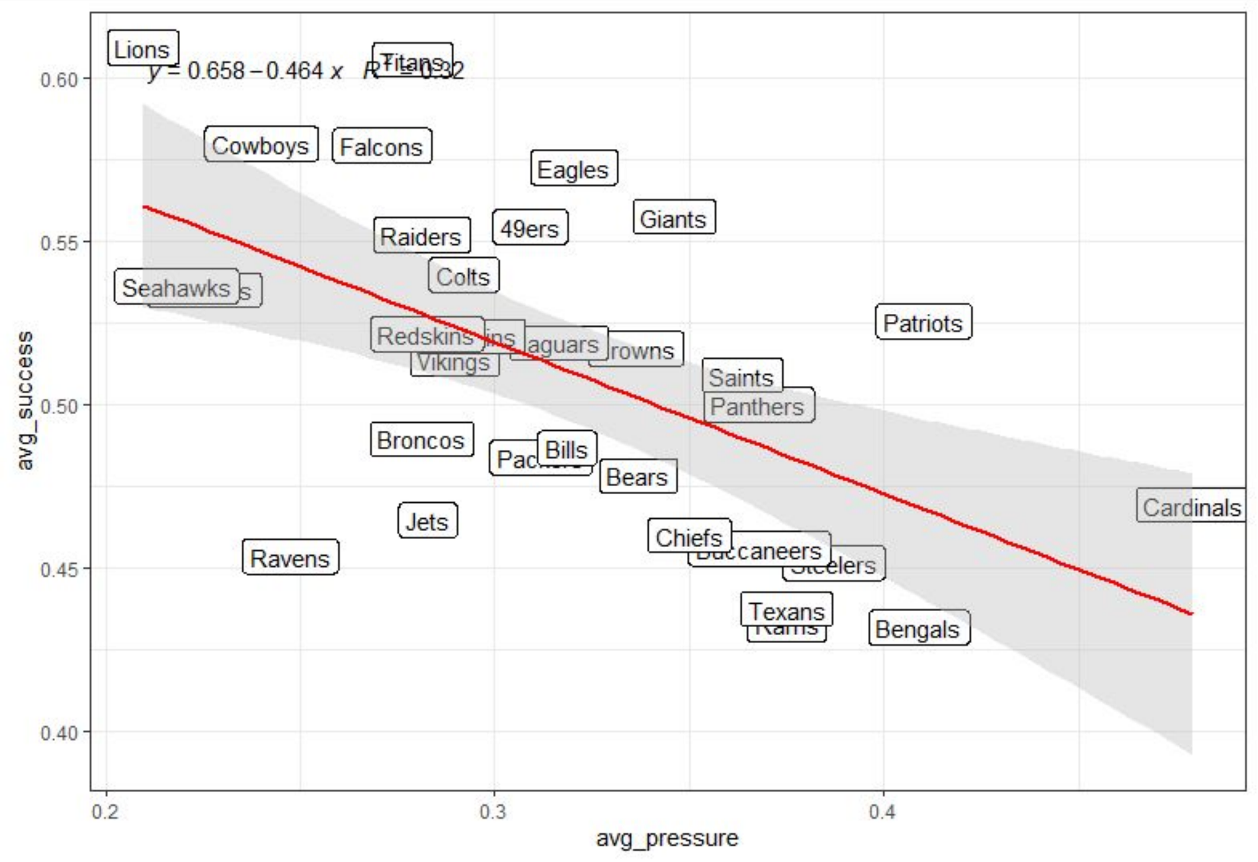
I took was to find what *defensive* within the data correlated to play success, for rushes and passes.

a. Correlation between success rate of pass plays and EPA (R-squared 0.59). This is also true for rush plays (R-squared 0.63)
b. The average pressure per play a defense generated is weakly correlated to their opposing offensives pass success rate, but not correlated to EPA.

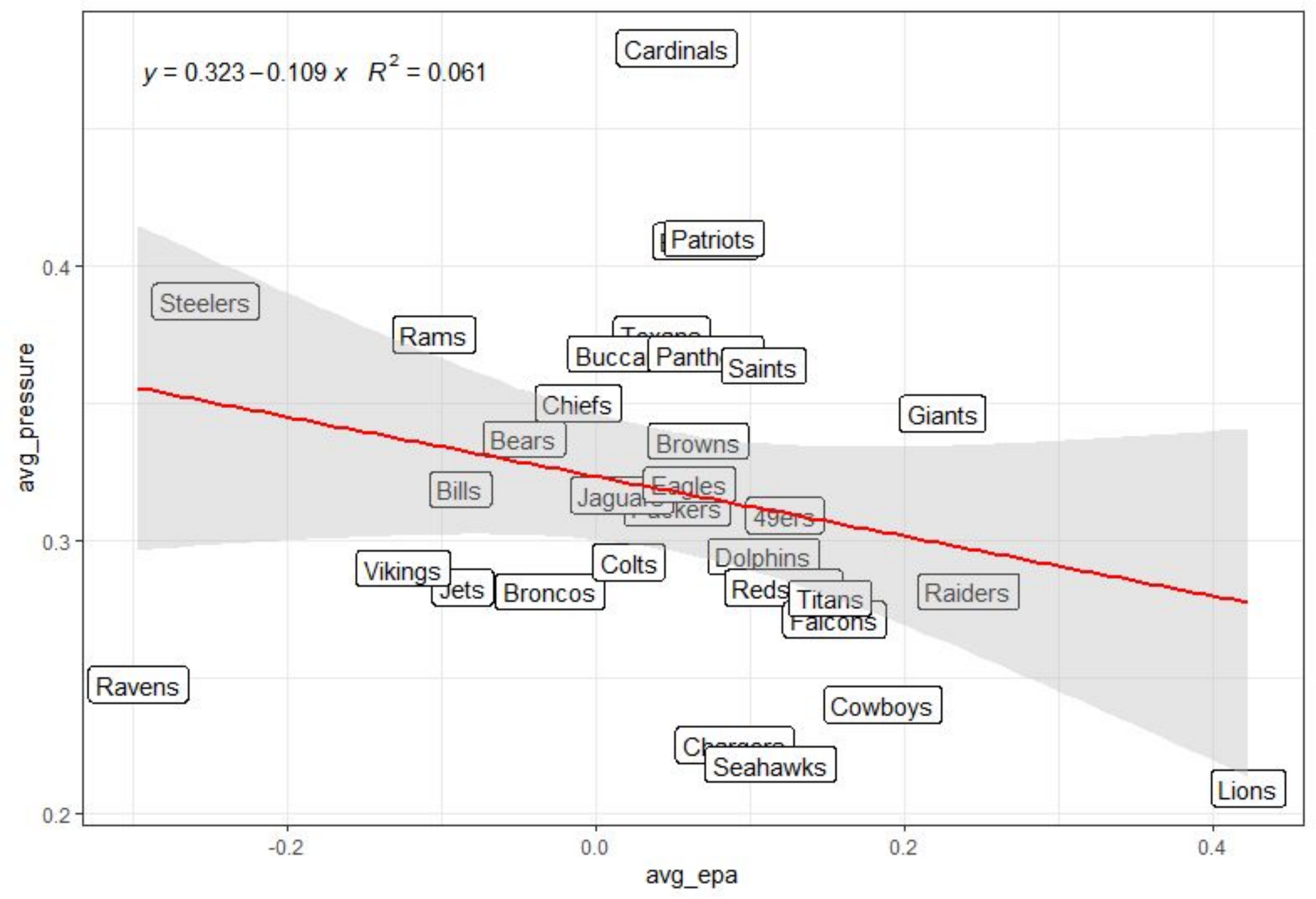
Assumption: By using success rate, I can be more flexible with the plays I include because success rate accounts for down-and-distance, and field position. The sample size is limited, and I want to use as much of it as possible. Additionally, success rate is not as volatile as EPA to 1st downs, turnovers, sacks, and touchdowns.



Correlation between success rate of pass plays and EPA (R-squared 0.59).



The average pressure per play a defense generated is weakly correlated to their opposing offensives pass success rate, but not correlated to EPA.



The average pressure per play a defense generated is weakly correlated to their opposing offensives pass success rate, but not correlated to EPA.

* = P < 0.05

Findings

Next, I looked at the data by play (excluding all individual player data). The goal of this analysis was to evaluate what was important on a play by play level to play success. Here are my key findings:

a. Pass plays with pressure had a success rate of ~ -23.3% less than pass plays without pressure*
b. If a team didn't use the intended rush gap on a run play, they had a success rate that was 4.3% less*
c. While blitzing didn't lead to a difference in success rate or EPA, it did result in a ~11% increase in pressure rate*.

Findings

Next, I looked at the data by position, including OnFieldPosition, RosterPosition, SideofBall, and TechniqueName, and rushing data (e.g. RunDirection)

a. There were some significant ($P < 0.05$) findings for OnFieldPosition (LB vs. DL), ,but these are due to their technique and will be discussed then.
b. There were some significant ($P < 0.05$) findings for RosterPosition, but some of these factors are due to their technique and will be discussed then.
c. When evaluating the “SideOfBall” the effect as whole is significant, however this is **only** driven by the “Outside” position for LB and DEs where $P \sim 0.0006143$ ($R-L \sim -2.3\%$).

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

Talent Distribution

Question 2 & 3

Limitations and Future Analysis

Findings

Next, looked
OnFieldPosi
TechniqueN

a. The
OnF que
and

b. The sition,
but so will be
discussed then.

c. When evaluating the “SideOfBall” the effect as whole is significant, however this is **only** driven by the “Outside” position for LB and DEs where P ~ 0.0006143 (R-L ~ -2.3%).

This is important because for in future analysis I will assume that “left” and “right” are equal when analyzing techniques and position on the line. *This is valid assumption for all techniques and roster positions with exception to LBs and DEs when they are in the “outside” technique*.

Essentially, I am assuming the gap between R-L exists for LBs and DEs is driven by **player talent**, not by technique or location.

Findings

By position, continued…

I couldn’t find any effects by player position on rush plays. The results were significant between what the offense chose to do (run direction). However, I analyzed each run direction type (e.g. A Gap) vs. the outcomes of the play and found no correlation or significance to players position or their ability to influence the intended gap OR make a tackle.

As a result, this analysis focuses on pass plays.

True Pressure Rate

Finally, I looked at the player level. Since I am focusing on the rate at which a player successfully rushes the passer, I calculated a new category:

True Pressure Rate = Rate of pressure the player generates when they are rushing (IsRushing = 1).

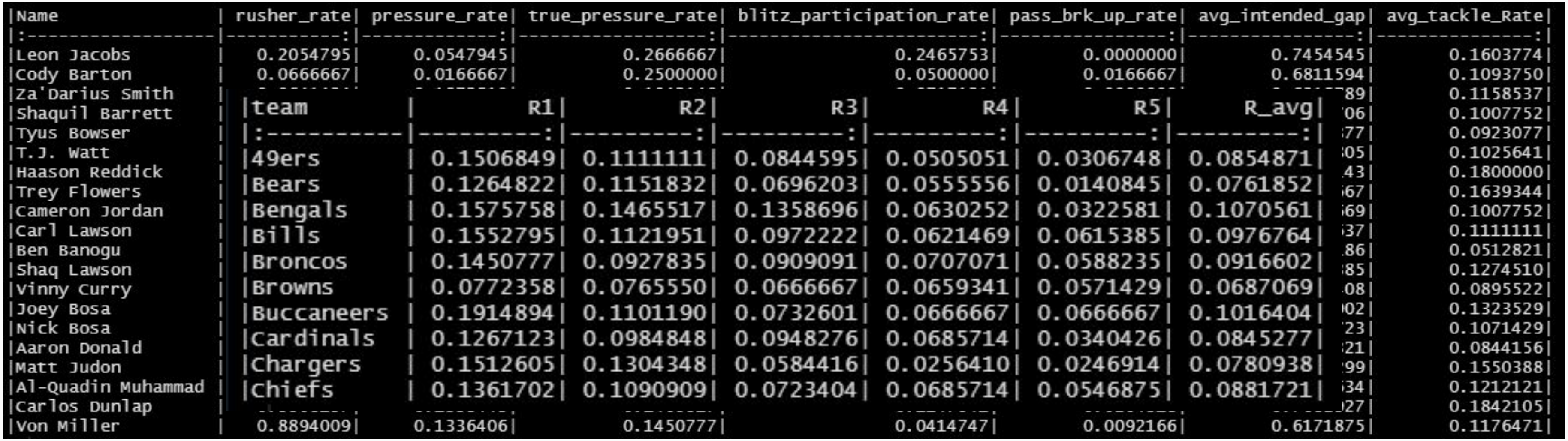
Using true pressure rate I looked the team level to try and identify the **cost of pressure**.

True Pressure Rate

Finally, I looked at the player level. Since I am focusing on the rate at which a player successfully rushes the passer, I calculated a new category:

True Pressure Rate = Rate of pressure the player generates when they are rushing (IsRushing = 1).

Using true pressure rate I looked the team level to try and identify the **cost of pressure**.



```
|Name                | rusher_rate| pressure_rate| true_pressure_rate| blitz_participation_rate| pass_brk_up_rate| avg_intended_gap| avg_tackle_Rate|
|:-------------------|-----------:|-------------:|------------------:|------------------------:|----------------:|----------------:|---------------:|
|Leon Jacobs         |   0.2054795|     0.0547945|          0.2666667|                0.2465753|        0.0000000|        0.7454545|       0.1603774|
|Cody Barton         |   0.0666667|     0.0166667|          0.2500000|                0.0500000|        0.0166667|        0.6811594|       0.1093750|
|Za'Darius Smith     |                                                                                                                  0.1158537|
|Shaquil Barrett     |                                                                                                                  0.1007752|
|Tyus Bowser         |                                                                                                                  0.0923077|
|T.J. Watt           |                                                                                                                  0.1025641|
|Haason Reddick      |                                                                                                                  0.1800000|
|Trey Flowers        |                                                                                                                  0.1639344|
|Cameron Jordan      |                                                                                                                  0.1007752|
|Carl Lawson         |                                                                                                                  0.1111111|
|Ben Banogu          |                                                                                                                  0.0512821|
|Shaq Lawson         |                                                                                                                  0.1274510|
|Vinny Curry         |                                                                                                                  0.0895522|
|Joey Bosa           |                                                                                                                  0.1323529|
|Nick Bosa           |                                                                                                                  0.1071429|
|Aaron Donald        |                                                                                                                  0.0844156|
|Matt Judon          |                                                                                                                  0.1550388|
|Al-Quadin Muhammad  |                                                                                                                  0.1212121|
|Carlos Dunlap       |                                                                                                                  0.1842105|
|Von Miller          |   0.8894009|     0.1336406|          0.1450777|                0.0414747|        0.0092166|        0.6171875|       0.1176471|
```

```
|team       |        R1|        R2|        R3|        R4|        R5|     R_avg|
|:----------|---------:|---------:|---------:|---------:|---------:|---------:|
|49ers      | 0.1506849| 0.1111111| 0.0844595| 0.0505051| 0.0306748| 0.0854871|
|Bears      | 0.1264822| 0.1151832| 0.0696203| 0.0555556| 0.0140845| 0.0761852|
|Bengals    | 0.1575758| 0.1465517| 0.1358696| 0.0630252| 0.0322581| 0.1070561|
|Bills      | 0.1552795| 0.1121951| 0.0972222| 0.0621469| 0.0615385| 0.0976764|
|Broncos    | 0.1450777| 0.0927835| 0.0909091| 0.0707071| 0.0588235| 0.0916602|
|Browns     | 0.0772358| 0.0765550| 0.0666667| 0.0659341| 0.0571429| 0.0687069|
|Buccaneers | 0.1914894| 0.1101190| 0.0732601| 0.0666667| 0.0666667| 0.1016404|
|Cardinals  | 0.1267123| 0.0984848| 0.0948276| 0.0685714| 0.0340426| 0.0845277|
|Chargers   | 0.1512605| 0.1304348| 0.0584416| 0.0256410| 0.0246914| 0.0780938|
|Chiefs     | 0.1361702| 0.1090909| 0.0723404| 0.0685714| 0.0546875| 0.0881721|
```

Example players with +100 plays registers, sorted by true pressure rate.

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

Talent Distribution

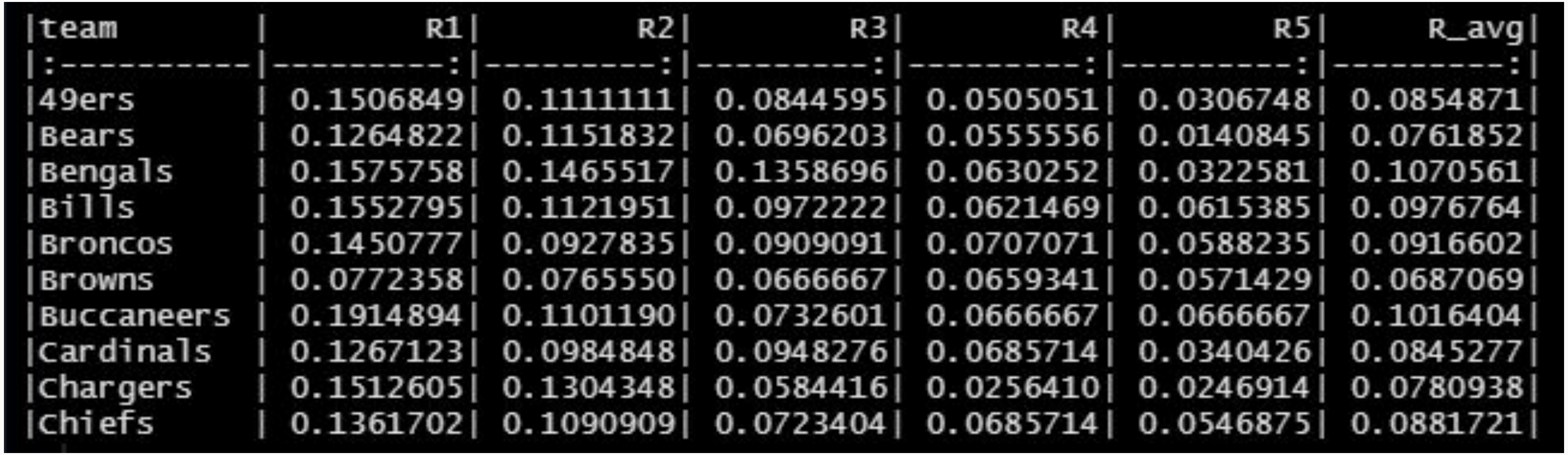
Question 2 & 3

Limitations and Future Analysis

True Pressure Rate

To evaluate the cost of pressure ~ or what correlates to a teams pressure rate from a player level I took the 5 most common players listed by team and sorted them by their true pressure rating.

*I also did this for the most common rushers and best rushers with a play minimum. The results hold across the samples.



```
|team       |        R1|        R2|        R3|        R4|        R5|     R_avg|
|:----------|---------:|---------:|---------:|---------:|---------:|---------:|
|49ers      | 0.1506849| 0.1111111| 0.0844595| 0.0505051| 0.0306748| 0.0854871|
|Bears      | 0.1264822| 0.1151832| 0.0696203| 0.0555556| 0.0140845| 0.0761852|
|Bengals    | 0.1575758| 0.1465517| 0.1358696| 0.0630252| 0.0322581| 0.1070561|
|Bills      | 0.1552795| 0.1121951| 0.0972222| 0.0621469| 0.0615385| 0.0976764|
|Broncos    | 0.1450777| 0.0927835| 0.0909091| 0.0707071| 0.0588235| 0.0916602|
|Browns     | 0.0772358| 0.0765550| 0.0666667| 0.0659341| 0.0571429| 0.0687069|
|Buccaneers | 0.1914894| 0.1101190| 0.0732601| 0.0666667| 0.0666667| 0.1016404|
|Cardinals  | 0.1267123| 0.0984848| 0.0948276| 0.0685714| 0.0340426| 0.0845277|
|Chargers   | 0.1512605| 0.1304348| 0.0584416| 0.0256410| 0.0246914| 0.0780938|
|Chiefs     | 0.1361702| 0.1090909| 0.0723404| 0.0685714| 0.0546875| 0.0881721|
```

Example teams with their 5 “best” rushers listed with their true pressure rates. R1 = highest true pressure rate.

True Pressure Rate

To evaluate the cost of pressure ~ or what correlates to a teams pressure rate from a player level I took the 5 most common players listed by team and sorted them by their true pressure rating. Here are how they correlated to their teams pressure rate:

1. R1 (best rusher to adj r-sq ~ 0.2855
2. R2 adj r-sq ~ 0.2539
3. R3 adj r-sq ~ 0.2139
4. R4 adj r-sq ~ 0.2045
5. R5 adj r-sq ~ -0.01068

*I also did this for the most common rushers and best rushers with a play minimum. The results generally hold across the samples.

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

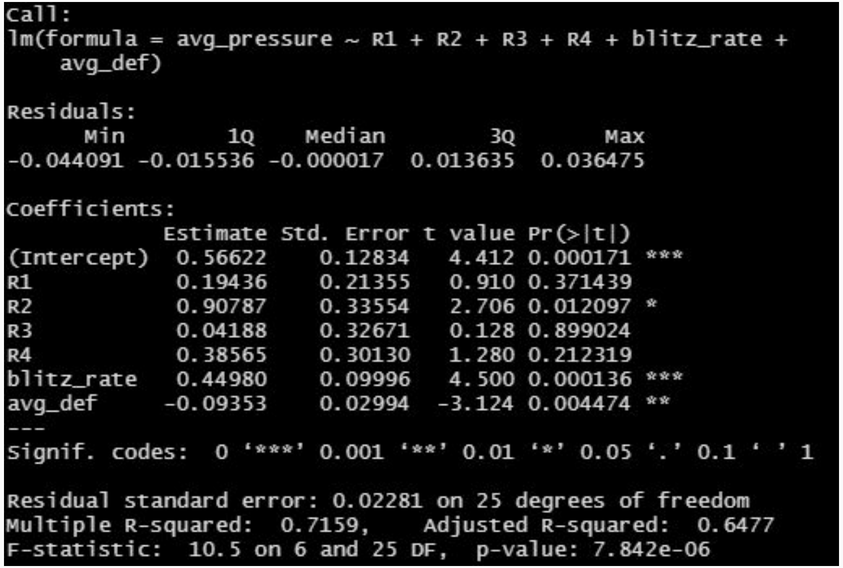
Talent Distribution

Question 2 & 3

Limitations and Future Analysis

True Pressure Rate

Additionally, by including the top 4 players, a teams blitz rate, and the average defenders listed, there is a mild-strong correlation to that teams pressure.



```
Call:
lm(formula = avg_pressure ~ R1 + R2 + R3 + R4 + blitz_rate +
    avg_def)

Residuals:
      Min        1Q    Median        3Q       Max
-0.044091 -0.015536 -0.000017  0.013635  0.036475

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.56622    0.12834   4.412 0.000171 ***
R1           0.19436    0.21355   0.910 0.371439
R2           0.90787    0.33554   2.706 0.012097 *
R3           0.04188    0.32671   0.128 0.899024
R4           0.38565    0.30130   1.280 0.212319
blitz_rate   0.44980    0.09996   4.500 0.000136 ***
avg_def     -0.09353    0.02994  -3.124 0.004474 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02281 on 25 degrees of freedom
Multiple R-squared:  0.7159,    Adjusted R-squared:  0.6477
F-statistic:  10.5 on 6 and 25 DF,  p-value: 7.842e-06
```

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

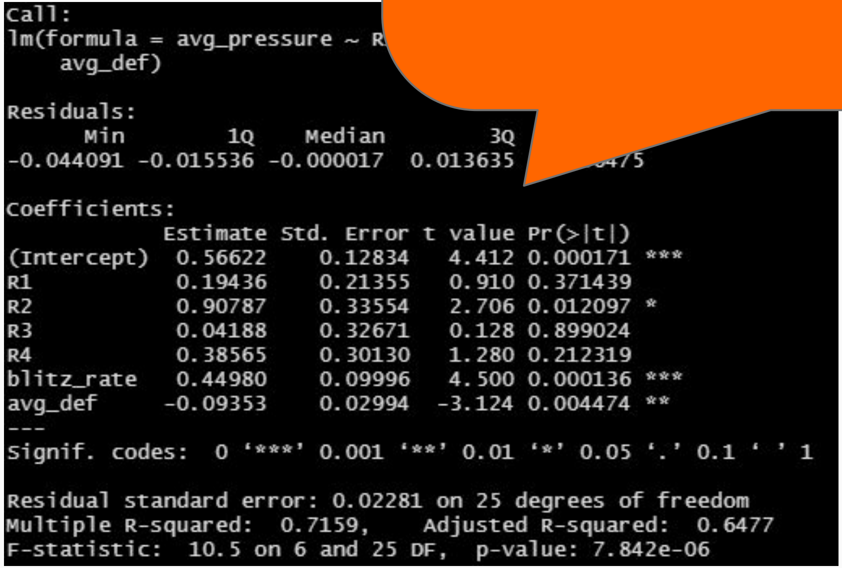
Talent Distribution

Question 2 & 3

Limitations and Future Analysis

True Press

Additionally, by inc
rate, and the avera
correlation to that



```
Call:
lm(formula = avg_pressure ~ R
    avg_def)

Residuals:
      Min        1Q    Median        3Q
-0.044091 -0.015536 -0.000017  0.013635      0475

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.56622    0.12834   4.412 0.000171 ***
R1           0.19436    0.21355   0.910 0.371439
R2           0.90787    0.33554   2.706 0.012097 *
R3           0.04188    0.32671   0.128 0.899024
R4           0.38565    0.30130   1.280 0.212319
blitz_rate   0.44980    0.09996   4.500 0.000136 ***
avg_def     -0.09353    0.02994  -3.124 0.004474 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02281 on 25 degrees of freedom
Multiple R-squared:  0.7159,    Adjusted R-squared:  0.6477
F-statistic:  10.5 on 6 and 25 DF,  p-value: 7.842e-06
```

The findings here mimic the assumption that a team that is better at rushing (better rushers) and blitzing generates more pressure.

Interestingly, the avg number of defenders on the front is negatively correlated to pressure. This is even more interesting when we look at the correlation to success in the next slide.

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

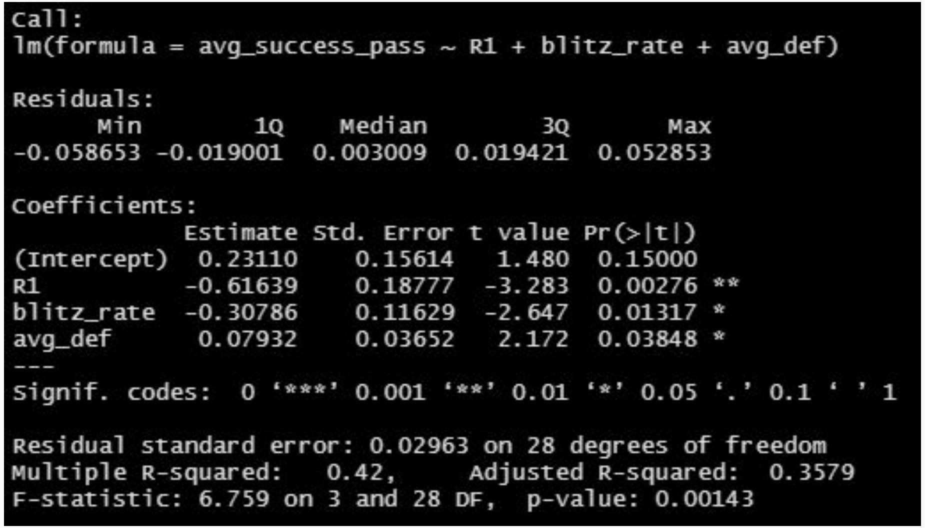
Talent Distribution

Question 2 & 3

Limitations and Future Analysis

True Pressure Rate

Looking at pass play success, we find a weak-mild correlation (adj. R-sq ~ 0.35) if we look at the best rusher, and a teams blitz and avg defenders up front.



```
Call:
lm(formula = avg_success_pass ~ R1 + blitz_rate + avg_def)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.058653 -0.019001  0.003009  0.019421  0.052853 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)  0.23110    0.15614   1.480  0.15000   
R1          -0.61639    0.18777  -3.283  0.00276 **
blitz_rate  -0.30786    0.11629  -2.647  0.01317 * 
avg_def      0.07932    0.03652   2.172  0.03848 * 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02963 on 28 degrees of freedom
Multiple R-squared:   0.42,	Adjusted R-squared:  0.3579 
F-statistic: 6.759 on 3 and 28 DF,  p-value: 0.00143
```

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

Talent Distribution

Question 2 & 3

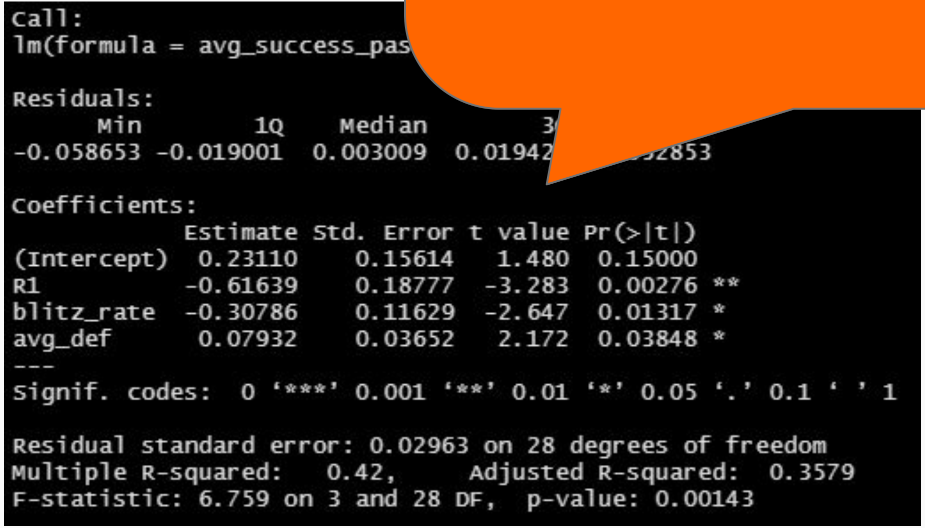
Limitations and Future Analysis

True Press

Looking at pass pl
correlation (adj. R-
and a teams blitz

Here, more average defenders on the line correlate to less success, aligning with the assumption that good defenses can get pressure with a fewer number of players (e.g. more players back to defend the pass).

Blitzing seems to act as a sort of “force multiplier”



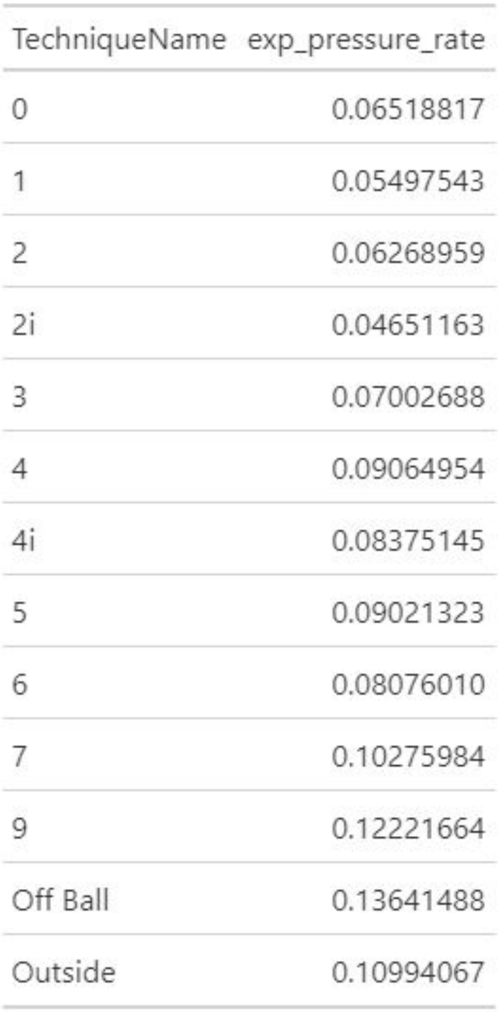
```
Call:
lm(formula = avg_success_pas

Residuals:
      Min         1Q    Median        3
-0.058653 -0.019001  0.003009  0.01942          2853

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.23110    0.15614   1.480  0.15000
R1          -0.61639    0.18777  -3.283  0.00276 **
blitz_rate  -0.30786    0.11629  -2.647  0.01317 *
avg_def      0.07932    0.03652   2.172  0.03848 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02963 on 28 degrees of freedom
Multiple R-squared:   0.42,     Adjusted R-squared:  0.3579
F-statistic: 6.759 on 3 and 28 DF,  p-value: 0.00143
```

Positional Value

Since the pressure rate by players (especially whose who play or rush frequently) is correlated to their teams pressure rate and success rate against the pass, I assume that this is a good proxy for value in the absence of other evidence.

This true pressure rate can be described as an “expected pressure rate”.

This rate does not adjust for the roster position DL, DE, or LB.

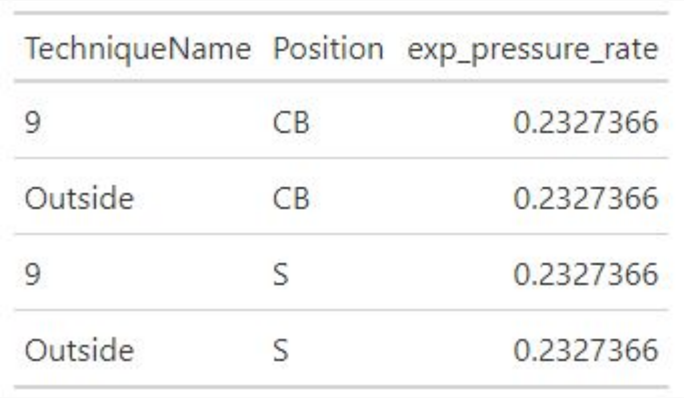


TechniqueName	exp_pressure_rate
0	0.06518817
1	0.05497543
2	0.06268959
2i	0.04651163
3	0.07002688
4	0.09064954
4i	0.08375145
5	0.09021323
6	0.08076010
7	0.10275984
9	0.12221664
Off Ball	0.13641488
Outside	0.10994067

Key Findings
Process & Assumptions
General Findings
True Pressure Rate
Positional Value
Question 1
Player Talent
Talent Distribution
Question 2 & 3
Limitations and Future Analysis

Positional Value

In this analysis, corners and safeties are separated and pooled because they predominantly rush from the outside and are significantly different than other other line positions when looking at rush rate, blitz participation rate, and true pressure rate. Due to their similarity and sample size, they are treated equally in all conditions listed.

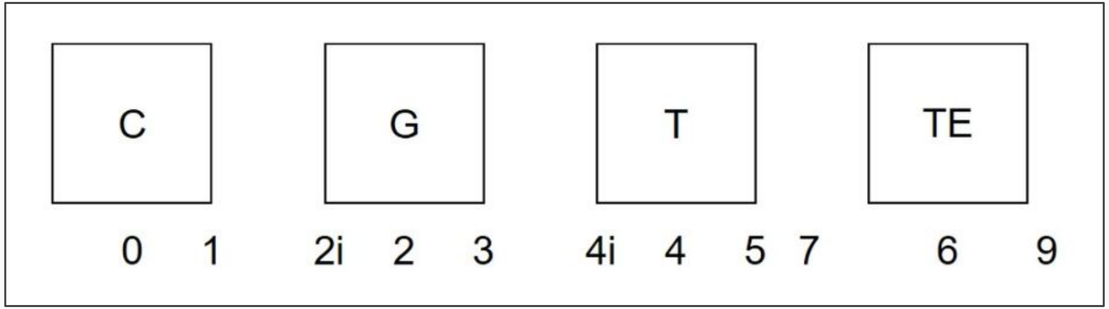


TechniqueName	Position	exp_pressure_rate
9	CB	0.2327366
Outside	CB	0.2327366
9	S	0.2327366
Outside	S	0.2327366

Question 1 Answer

Which is the most valuable defensive line position, as you define them?

The most valuable defensive line positions are those on the exterior (further outside, regardless of the side of ball) because they are more effective at generating pressure when they are asked to rush ~ or their “true pressure rate”.



Value increases with “expected pressure rate” →

Positional Value

Looking at the RosterPositions of the players, their positional value (as derived from their expected pressure rate is as follows):

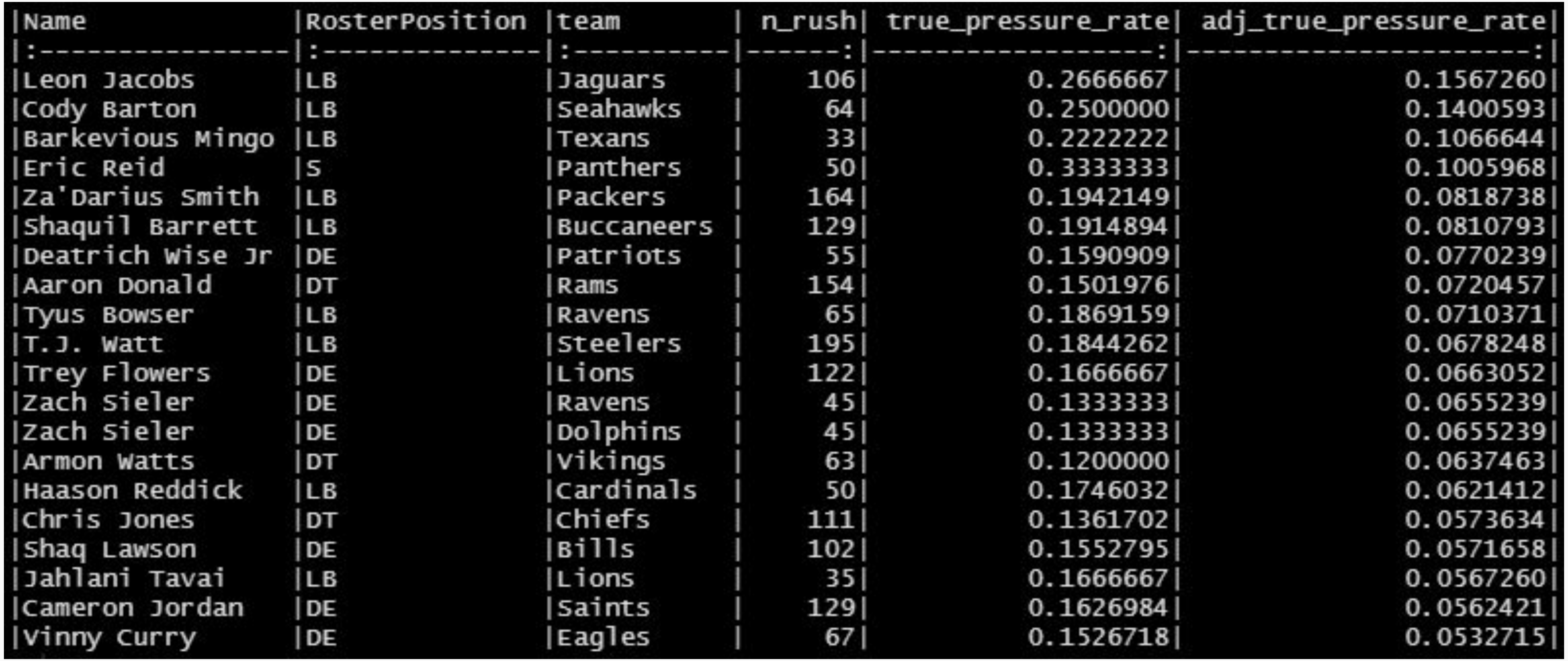
1. LB (highest value)
2. DE
3. DT (lowest value)

However, this difference is driven by what technique they play. This is also true for if they had a hand on the ground or not.

Player Talent

Next, to evaluate the skill of each player, I looked at each play they were were listed and deducted the expected pressure rate from their actual pressure rate on plays that they were rushing. This is a players:

Adjusted True Pressure Rate = the rate at which a player generates pressure on plays that they are rushing, adjusted for the average pressure that that player would be expected to generate given their technique.



```
|Name             |RosterPosition |team       | n_rush| true_pressure_rate| adj_true_pressure_rate|
|:----------------|:--------------|:----------|------:|------------------:|----------------------:|
|Leon Jacobs      |LB             |Jaguars    |    106|          0.2666667|              0.1567260|
|Cody Barton      |LB             |Seahawks   |     64|          0.2500000|              0.1400593|
|Barkevious Mingo |LB             |Texans     |     33|          0.2222222|              0.1066644|
|Eric Reid        |S              |Panthers   |     50|          0.3333333|              0.1005968|
|Za'Darius Smith  |LB             |Packers    |    164|          0.1942149|              0.0818738|
|Shaquil Barrett  |LB             |Buccaneers |    129|          0.1914894|              0.0810793|
|Deatrich Wise Jr |DE             |Patriots   |     55|          0.1590909|              0.0770239|
|Aaron Donald     |DT             |Rams       |    154|          0.1501976|              0.0720457|
|Tyus Bowser      |LB             |Ravens     |     65|          0.1869159|              0.0710371|
|T.J. Watt        |LB             |Steelers   |    195|          0.1844262|              0.0678248|
|Trey Flowers     |DE             |Lions      |    122|          0.1666667|              0.0663052|
|Zach Sieler      |DE             |Ravens     |     45|          0.1333333|              0.0655239|
|Zach Sieler      |DE             |Dolphins   |     45|          0.1333333|              0.0655239|
|Armon Watts      |DT             |Vikings    |     63|          0.1200000|              0.0637463|
|Haason Reddick   |LB             |Cardinals  |     50|          0.1746032|              0.0621412|
|Chris Jones      |DT             |Chiefs     |    111|          0.1361702|              0.0573634|
|Shaq Lawson      |DE             |Bills      |    102|          0.1552795|              0.0571658|
|Jahlani Tavai    |LB             |Lions      |     35|          0.1666667|              0.0567260|
|Cameron Jordan   |DE             |Saints     |    129|          0.1626984|              0.0562421|
|Vinny Curry      |DE             |Eagles     |     67|          0.1526718|              0.0532715|
```

Example of the players with the highest adjusted true pressure rate given a play minimum (30)

Key Findings
Process & Assumptions
General Findings
True Pressure Rate
Positional Value
    Question 1
Player Talent
Talent Distribution
    Question 2 & 3
Limitations and Future Analysis

Player Talent

Testing the results:

True Adjusted Pressure Rate is not significantly different between any 2 categories for for OnFieldPosition or RosterPosition, meaning it is correctly adjusting across the roster positions.

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

Talent Distribution

Question 2 & 3

Limitations and Future Analysis

Talent Distribution

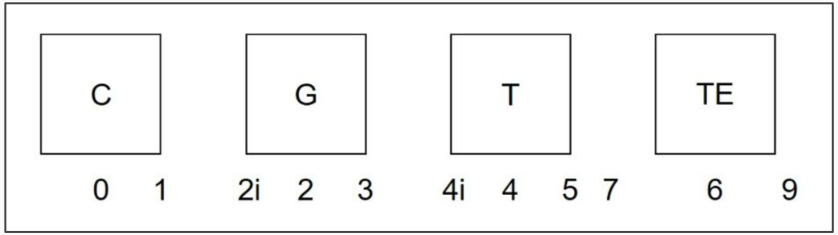
To evaluate talent distribution, I grouped the techniques based on expected pressure rate.

0, 1, 2i, 2 = Interior

3, 4i, 4, 5 = EDGE

7, 6, 9, Outside = EDGE Plus

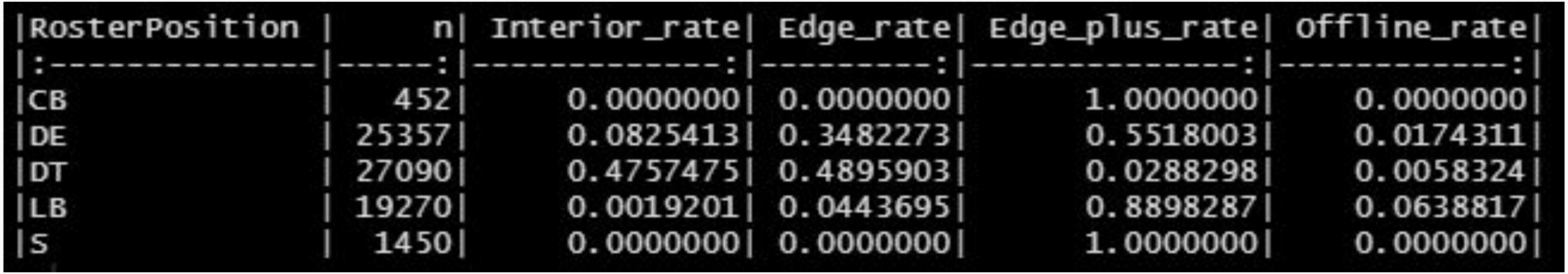
Off ball = Off ball



Talent Distribution

This is a grouping based on expected pressure rate, and each group is significantly different from each other (The groups are different).

Additionally, as expected the True Adjusted Pressure Rate is not significantly different between any 2 categories, meaning it is correctly adjusting across these positions.



```
|RosterPosition |     n| Interior_rate| Edge_rate| Edge_plus_rate| Offline_rate|
|:--------------|-----:|-------------:|---------:|--------------:|------------:|
|CB             |   452|     0.0000000| 0.0000000|      1.0000000|    0.0000000|
|DE             | 25357|     0.0825413| 0.3482273|      0.5518003|    0.0174311|
|DT             | 27090|     0.4757475| 0.4895903|      0.0288298|    0.0058324|
|LB             | 19270|     0.0019201| 0.0443695|      0.8898287|    0.0638817|
|S              |  1450|     0.0000000| 0.0000000|      1.0000000|    0.0000000|
```

Example at which each RosterPosition was in one of the new positions.

Key Findings
Process & Assumptions
General Findings
True Pressure Rate
Positional Value
    Question 1
Player Talent
Talent Distribution
    Question 2 & 3
Limitations and Future Analysis

Talent Distribution

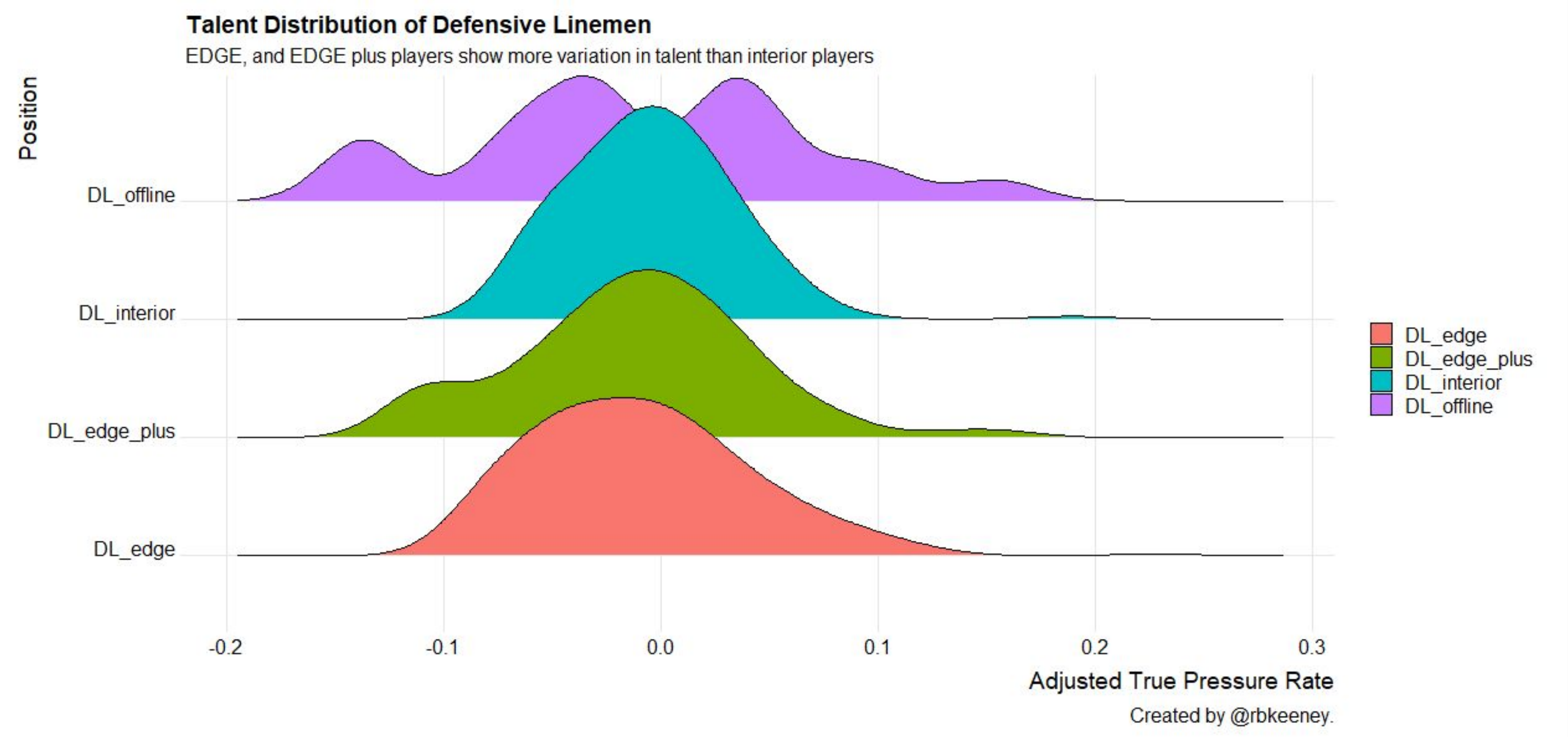
To calculate the talent distribution I grouped player by their new position group for each play and calculated the distribution for each position group.

Since each group is centered about 0 (as expected, because the average player should see +/- 0 expected pressure rate, I evaluated them by their standard deviation.

Interior, sd = 0.0371 (least variance, lowest differentiation in talent)

EDGE = EDGE Plus, sd = 0.050 (most variance, lowest differentiation in talent)

This assumes we should exclude the Off Line, sd = 0.077, sample.

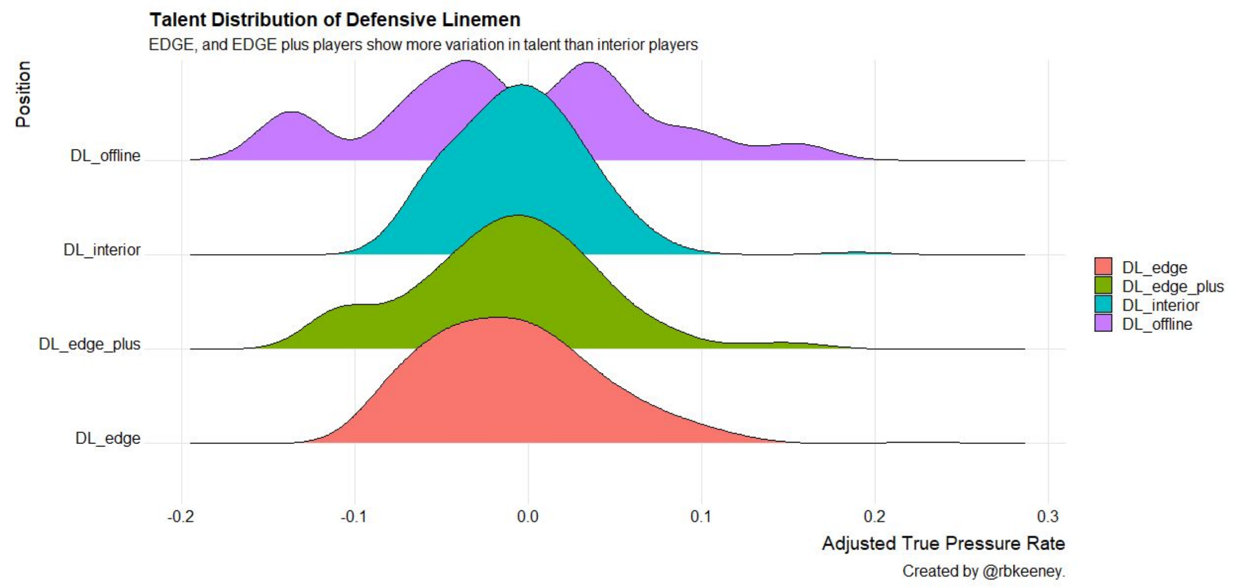


Interior players show less variation in talent that EDGE and EDGE plus type players.

Question 2 Answer

What is the nature of the distribution of talent between the defensive line positions, as you define them?

Interior players show less variation in talent that EDGE and EDGE plus type players when comparing their adjusted true pressure rates.



Talent Distribution

To calculate the talent distribution I grouped player by their new position group for each play and calculated the distribution for each position group.

Since each group is centered about 0 (as expected, because the average player should see +/- 0 expected pressure rate, I evaluated them by their standard deviation.

Interior, sd = 0.0371 (least variance, lowest differentiation in talent)

EDGE = EDGE Plus, sd = 0.050 (most variance, lowest differentiation in talent)

This assumes we should exclude the Off Line, sd = 0.077, sample.

Question 3 Answer

Not all situations are created equal. In which in-game or roster construction scenarios would the answer to Question 1 change?

In Game situations

Since the adjusted pressure rates per the groups (interior, edge, edge-plus) do not show any significant differences between the blitz and shotgun plays I am comfortable with my findings in relation to these aspects.

This may be because DBs (from outside) have a high rate of expected pressure and I separated them in my analysis.

Question 3 Answer

Not all situations are created equal. In which in-game or roster construction scenarios would the answer to Question 1 change?

Roster construction

Looking at pass play success, we find a weak-mild correlation (adj. R-sq ~ 0.35) if we look at the **best** rusher, and a teams blitz and avg defenders up front.

In fact, the single best pass rusher (in terms of true pressure rate) had a adj. R-sq ~ 0.1717 to their defenses pass success rate. Not accounting for CB play, they may be the single most important player on a defense. This player doesn't not have to play on the outside (the more valuable position), but it is suggested because it seems to be easier to generate pressure there.

Additionally, there is no correlation between a teams best rusher (in terms of true pressure rate) and their blitz rate. Since blitz rate seems to act as a force multiplier, this seems like a possible area for teams with the best rushing talent to exploit, especially with S or CB rushes outside as they have a high rate of success at generating pressure.

Key Findings

Process & Assumptions

General Findings

True Pressure Rate

Positional Value

Question 1

Player Talent

Talent Distribution

Question 2 & 3

Limitations and Future Analysis

Limitations and Future Analysis

- With only ~ only ½ season, I didn’t feel confident to data to make prediction on some unique down-and-distance and expected-win-% scenarios. I chose to keep my analysis including all plays. Additionally, with more data I could have looked at how well predictive certain stats were (e.g. compare games 1-8 v. 9-16)
- Right now I am “blind to the offensive personnel and alignment. I assume defensive personnel is dictated by the offense (shown some by the differences in shotgun plays).If we had offensive personal and alignment I could investigate the effects of OL talent, and how the different alignments affected the defense.

Limitations and Future Analysis

- With additional information on defenders in the box (for better run defense information), I could model how second level players like linebackers affected the outcomes of play (similar to Josh Hermsmeyer's analysis on field position and defenders in the box to rushing yards.
- With player tracking, or additional player attributes like double teams, time to throw, rush "win" rates, there are many other factors of line play that could be e.g. (I don't know if a DT often takes up 2x offensive linemen, giving other rushers with an advantage).

Fin.

Thanks, Ryan